

REMARKS

Applicants have now had an opportunity to consider the Examiner's Final Office Action of November 14, 2002. Claims 1-20 remain in the application. Reexamination and reconsideration of the application is respectfully requested.

THE OFFICE ACTION

Claims 1, 5-7, 9, 15 and 20 are rejected under 35 U.S.C. 102(a) as being anticipated by U.S. Patent No. 5,885,083 to Ferrell.

Claims 2-4 16-18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Ferrell in view of US Patent No. 5,878,396 to Henton.

Claims 8 and 11-14 are rejected under 35 U.S.C. 103(a) as being unpatentable over Ferrell in view of US Patent No. 5,920,838 to Mostow et al.

Claims 10 and 19 are rejected under 35 U.S.C. §103(a) as being unpatentable over Ferrell in view of Henton and Adams, Jr. et al.

INTERVIEW SUMMARY

Prior to the filing of the present Continued Prosecution Application (CPA) on October 24, 2002, the applicants' attorney and the Examiner conducted a telephone interview regarding the Final Action in the parent application. During the interview, the applicants' attorney asserted distinctions of the claimed invention over the prior art. The Examiner did not agree with these distinctions. Accordingly, no agreement as to allowability of claims was reached.

After receiving the present Final Examiner's Action, which is the first action on this application, the applicants' attorney telephoned the Examiner. The Examiner indicated that if the response to the Final Office Action placed the claims in condition for allowance, then they would be allowed. The applicants submit that the present response does so.

Accordingly, the claims should be allowed. In the event that the Examiner has any questions or concerns, he is invited to telephone the undersigned attorney.

THE CLAIMS DISTINGUISH OVER THE TEACHINGS OF THE CITED PATENTS

The Examiner rejected claims 1, 5-7, 9, 15 and 20 under 35 U.S.C. §102(a) as being anticipated by Ferrell. However, for the reasons set forth below, the claims are not anticipated by this cited patent.

Specifically, Ferrell relates to a system designed to teach jargon to particular professions, for example, air traffic controllers. The contemplated focus of the development relates to time-critical situations, so the training is intended to allow the user to quickly learn to properly respond, e.g. by reflex. This patent relates to correctness of pronunciation and does not concern confidence measures regarding precision. In this regard, the Examiner (in the recent interview summary) indicates that Ferrell teaches a confidence measure. However, Ferrell is concerned only with correctness, not precision. Given the environment in which Ferrell is disclosed, correctness is a reasonable objective. In the present invention, however, the **precision** at which a user replicates audible speech is the goal. Precision is recited in all independent claims.

Moreover, in independent claims 1, 16, and 17, the precision is scored based on a comparison of utterances to either audible speech or a model. In Ferrell, the correctness is measured strictly based on predetermined criteria which is not specified. To further emphasize, claim 16 has been amended to define the model as being one of a predictive, phoneme, diphone or dynamically generated model. Therefore, at least these claimed features of the present application are not specifically disclosed in Ferrell. Moreover, because the present invention is not limited to measures based on predetermined criteria, the present invention allows for generation of uttered speech or models for comparison that can be generated "on the fly." Ferrell is apparently only concerned with predetermined words and criteria.

In addition, the independent claims have been amended to recite that each portion of the speech is relevant for the confidence measure -- not simply whether the word is "correct." Indeed, this allows for a system, such as that described in Figure 11, which scores speech as to portions of words or even letters. Such a feature is not anticipated by Ferrell.

Claims 2-4 and 16-18 were rejected under 35 U.S.C. §103 as being unpatentable over Ferrell in view of Henton. Apparently, Henton is cited for its teachings of synthetic speech with an animated face. However, even if Henton were somehow combinable with Ferrell, Henton does not overcome the deficiencies of Ferrell. Therefore, a suggested combination of Ferrell and Henton does not render the claims 2-4 and 16-18 obvious.

Claims 8 and 11-14 were rejected under 35 U.S.C. §103 as being unpatentable over Ferrell in view of Mostow. Mostow is cited for its asserted teaching of supplying text and for its asserted teaching of referencing tables. Again, however, even if Mostow was somehow combinable with Ferrell, the citation to Mostow does not cure the deficiencies of Ferrell. Therefore, the claims (8 and 11-14) are not rendered obvious.

Claims 10 and 19 were rejected under 35 U.S.C. §103 as being unpatentable over Ferrell, Henton, and Adams, Jr. Adams, Jr. is apparently cited for its teaching of "replay" features. However, citation to Adams, Jr. does not cure the deficiencies of the Ferrell/Henton combination to render the claims obvious. Therefore, claims 10 and 19 are not rendered obvious.


CONCLUSION

In view of the foregoing, it is respectfully submitted that this application is now in condition for allowance. An early notice to that effect is therefore earnestly solicited.

Respectfully submitted,

FAY, SHARPE, FAGAN,
MINNICH & McKEE, LLP

Date: April 29, 2003



Joseph D. Dreher, Reg. No. 37,123
1100 Superior Avenue
Seventh Floor
Cleveland, Ohio 44114-2518
(216) 861-5582

Attachment: Version with Markings to Show Changes Made

VERSION WITH MARKINGS TO SHOW CHANGES MADE

IN THE CLAIMS

Please amend claims 1, 16 and 17 as follows:

1. (Twice Amended) A system for interactive language instruction comprising:
 - a first module configured to convert input text to audible speech in a selected language, the audible speech being patterned after a model;
 - a user interface configured to receive utterances spoken by a user in response to a prompt to replicate the audible speech; and,
 - a second module configured to recognize the utterances and provide feedback to the user, the feedback being comprised of a confidence measure reflecting a precision at which the user replicates each portion of the audible speech in the selected language based on a comparison of the utterances to one of the audible speech and the model.

16. (Twice Amended) A system comprising:
 - a first module configured to convert input text to audible speech in a selected language, the audible speech indicative of a model, wherein the model is one of a predictive model, a phoneme model, a diphone model, and a dynamically generated model;
 - a second module synchronized to the first module, the second module producing an animated image of a human face and head pronouncing the audible speech;
 - a user interface positioned to receive utterances spoken by a user in response to a prompt to replicate the audible speech; and,
 - a third module configured to recognize the utterances and provide feedback to the user, the feedback being comprised of at least one of a score, an icon and an audio segment reflecting a precision at which the user replicates each portion of the speech in the selected language based on a comparison of the utterances to one of the audible speech and the model.

17. (Twice Amended) A method for voice interactive language instruction comprising:
converting input text data to audible speech data;
generating audible speech comprising phonemes based on the audible speech data;
outputting the audible speech through an audio output device;
generating an animated image of a face and head pronouncing the audible speech;
synchronizing the audible speech and the video image;
prompting a user to replicate the audible speech;
recognizing utterances generated by the user in response to the prompting;
comparing the audible speech to the utterances; and,
providing feedback to the user based on the comparison, the feedback comprised
of at least one of a score, an icon and an audio segment reflecting a precision at which the user
replicates each portion of the audible speech.